University of Groningen

Peak power output in handcycling of individuals with a chronic spinal cord injury

HandbikeBattle Grp

Published in:
Disability and Rehabilitation

DOI:
10.1080/09638288.2018.1501097

Manuscript title: **Peak power output in handcycling of individuals with a chronic spinal cord injury: predictive modeling, validation and reference values**

Running head: Predictive modeling in handcycling

Article category: Research Paper

Ingrid Kouwijzer[1,2,3], Linda Valent[1], Rutger Osterthun[4,5,6], Lucas van der Woude[2,5], and Sonja de Groot[2,3] on behalf of the HandbikeBattle group*

Affiliations

[1] Research and Development, Heliomare Rehabilitation Center, Wijk aan Zee, The Netherlands

[2] University of Groningen, University Medical Center Groningen, Center for Human Movement Sciences, Groningen, The Netherlands

[3] Amsterdam Rehabilitation Research Center | Reade, Amsterdam, The Netherlands

[4] Jeroen Bosch Hospital, Tolbrug Rehabilitation Centre, 's-Hertogenbosch, The Netherlands

[5] University of Groningen, University Medical Center Groningen, Center for Rehabilitation, Groningen, The Netherlands

[6] Rijndam Rehabilitation Institute, Rotterdam, The Netherlands

Corresponding author: Ingrid Kouwijzer, Research and Development, Heliomare Rehabilitation Center, Relweg 51, 1949 EC Wijk aan Zee, The Netherlands. Phone: 0031 (0)251-288013, i.kouwijzer@heliomare.nl

Word count paper: 5569 words

Word count abstract: 200 words

Abstract

Purpose: To develop and validate predictive models for peak power output to provide guidelines for individualized handcycling graded exercise test protocols for people with spinal cord injury; and to define reference values.

Material and methods: Power output was measured in 128 handcyclists with spinal cord injury during a synchronous handcycling exercise test. 80% of the data was used to develop four linear regression models: two theoretical and two statistical models with peak power output (in W and W/kg) as dependent variable. The other 20% of the data was used to determine agreement between predicted versus measured power output. Reference values were based on percentiles for the whole group.

Results: Lesion level, handcycling training hours and sex or body mass index were significant determinants of peak power output. Theoretical models (R^2=42%) were superior to statistical models (R^2=39% for power output in W, R^2=30% for power output in W/kg). The intraclass correlation coefficients varied between 0.35-0.60, depending on the model. Absolute agreement was low.

Conclusions: Both models and reference values provide insight in physical capacity of people with spinal cord injury in handcycling. However, due to the large part of unexplained variance and low absolute agreement, they should be used with caution.

Keywords

Arm ergometry, graded exercise test, physical capacity, normative values, post-rehabilitation

Introduction

Today synchronous handcycling has become a popular sport for wheelchair users (1). This is not surprising since handcycling is a relatively easy mode to cover large distances at a high speed compared to handrim wheelchair propulsion (1). Benefits of handcycling include its higher efficiency and lower strain compared to wheelchair propulsion, possibly reducing the risk of upper body overuse injuries (2–4). Moreover, it has been shown that handcycling can be a good way to improve physical capacity in, for example, individuals with a spinal cord injury (SCI) already early in rehabilitation (5). This is an important result, as the physical capacity in this population is generally low due to muscle paralysis and loss of sympathetic control under the lesion level, as well as a sedentary lifestyle (6–9). In previous studies, the benefits of an improvement in physical capacity for wheelchair users with a SCI have already been shown, such as a more favorable lipid profile (10), a higher life satisfaction (11,12) and a higher chance to return to work (13,14).

Abovementioned results are predominantly based on studies that focused on wheelchair capacity, which is different from handcycling, as demonstrated by the lower submaximal strain and higher peak power output (POpeak) during handcycling (3,4). Next to wheelchair ergometry, asynchronous arm ergometry is studied in individuals with SCI (15–18). However, several studies highlighted differences in physiological responses between the asynchronous and synchronous propulsion mode (15,19). For example, a higher net and gross efficiency, and a higher POpeak were found during asynchronous arm cranking compared to synchronous arm cranking (15,19). Therefore, results of these studies investigating asynchronous arm ergometry cannot be applied to the synchronous handcycling propulsion mode investigated in the present study. This emphasizes the importance of specificity in testing when studying submaximal and peak physiological responses.

In order to stimulate an improvement in physical capacity by means of handcycling in wheelchair users with SCI, the HandbikeBattle is organized as an annual event since 2013. The HandbikeBattle is an uphill handcycling mountain race in Austria in which currently 11 Dutch rehabilitation centers participate with approximately 6 participants each (20). All participants are chronic wheelchair users and relatively inexperienced handcyclists who train between 4 and 6 months prior to the event. Prior to participation, medical screening including a peak handcycle or synchronous arm crank aerobic exercise test (GXT) is obligatory. The GXT is part of the cardiopulmonary check-up and forms the basis for an individualized training guideline. When using a typical one-minute protocol and preferred GXT duration of 8 – 12 minutes (21), the anticipated POpeak (W) defines step size of the protocol. As many factors play a role in determining the potential physical capacity of these highly diverse individuals with SCI (9), it is hard to estimate each individual's POpeak prior to testing. As such it is difficult to select an optimal GXT protocol. It is, however, essential to select the right individualized protocol for an individual with SCI as the protocol itself affects actual peak performance (21–23). When the step size or ramp slope is too small or too large and, consecutively, test time is too long or too short, it will be unclear whether the "true" peak physical capacity is reached (21,22,24,25). Moreover, training guidelines based on these peak values will be non-optimal (21,22,24,25).

To select an optimal individual handcycling GXT protocol for individuals with SCI and, consecutively, improve the development of individualized training guidelines, a POpeak prediction model could be valuable. In such models, POpeak is estimated based on known participant characteristics. Moreover, development of a model could give a theoretical background in the underlying factors influencing physical capacity in individuals with SCI during handcycling and insight in which factors should be influenced to increase physical capacity. In addition to merely statistics-driven modeling, theory-driven statistical models could be useful to further clarify and explain the associations of underlying determinants with physical capacity for this specific mode of exercise.

Based on previous literature investigating wheelchair ergometry or asynchronous arm ergometry in individuals with SCI, several participant characteristics were identified to be of influence on POpeak. Sex, for example, showed to be an important characteristic, as women generally produce a lower POpeak than men (26), which might be explained by the smaller upper-body muscle mass (27). Moreover, lesion level and completeness are inversely related to POpeak (9,17,18,26,28–30). POpeak also declines with age (17,26,29,31) and increases with activity level (9,17,29,32,33). Time since injury (TSI) could be a determinant as physical capacity shows an increase in the first years after SCI (9,34,35) but thereafter seems to decrease (9,36) Janssen et al. (N=166) performed a statistical stepwise (forward) multiple regression analysis for POpeak in wheelchair ergometry and found lesion level, hours of sport, age, body mass, TSI and completeness to be significant determinants (with a cumulative explained variance (R^2) of 80%) (29). Simmons et al. (N=179) found functional classification, BMI and motor level of injury to be significant determinants for relative POpeak (W/kg) in (asynchronous) arm ergometry (cumulative R^2 of 57%) and motor level of injury, functional classification and sex for absolute POpeak (W) (cumulative R^2 of 57%), performing a forward multiple regression analysis (18). To date, in synchronous handcycling it is, however, still unknown which factors determine physical capacity. Moreover, previously described models have never been validated. Therefore, the validity of these models for use in clinical practice remains uncertain. Next to the missing knowledge about underlying factors influencing physical capacity in handcycling and uncertainty about the validity of predictive modeling, comparison to group level is lacking, as handcycling reference values for physical capacity for individuals with a SCI are scarce.

The aims of this study were, therefore:

1) To develop four predictive models: two theory-driven and two statistically-driven models for POpeak (W and W/kg) in a synchronous handcycling GXT for people with SCI.
2) To validate the four predictive models for POpeak.
3) To define reference values for absolute and relative POpeak and peak oxygen uptake (VO_2peak) in handcycling based on lesion level and sex.

Material and methods

Participants

Participants were retrospectively selected from the HandbikeBattle 2013, 2014, 2015, 2016 and 2017 cohorts. Every year was a unique cohort. Selection criteria for this study were having an SCI or spina bifida and the availability of comprehensive testing results. A total of 168 participants with SCI or spina bifida were selected. Forty participants were excluded due to missing data in either outcome variables or determinants. This led to 128 recreational handcyclists with SCI or spina bifida being included in this study. Participant characteristics are listed in table 1. The study was approved by the Local Ethical Committee of the Center for Human Movement Sciences, University Medical Center Groningen, the Netherlands (ECB/2012_12.04_l_rev/Ml). All participants voluntarily signed an informed consent form after they were given information about the testing procedures. The study was registered in the Dutch Trial Register (www.trialregister.nl, NTR6586).

Outcomes

In this cross-sectional study, participants underwent a medical screening including a medical history and a physical examination obtained by a physician. Moreover, all participants performed a GXT as part of the medical screening. As the GXT took place before the training period, participants were relatively untrained handcyclists. Depending on the rehabilitation center the pre-training GXT was performed with the use of an arm ergometer (Lode Angio, Groningen, The Netherlands) or a recumbent sport handcycle attached to the Tacx roller (Tacx, Terneuzen, The Netherlands) or Cyclus 2 ergometer (RBM elektronik-automation GmbH, Leipzig, Germany). Comparable peak physiological responses are to be expected between these ergometers (ICC 0.87 Lode vs Tacx, ICC 0.88 Lode vs Cyclus2) (37). All tests were performed in synchronous mode of cranking. A testing guideline and instructions were provided to the test assistants of all centers to make the tests as uniform as possible. Either a one-minute step protocol or continuous ramp protocol was used, depending on the preference and practice of the test assistant in the different rehabilitation centers. There was no systematic difference in VO_2peak and POpeak to be expected between these protocols (38). For the one-minute protocol, the test started at 20-30 W with increments of 5-15 W/min. For the ramp protocol, the test started at 0 W with a slope of 1 W / 12 sec (5 W/min), 1 W / 6 sec (10 W/min), 1 W / 4 sec (15 W/min) or 1 W / 3 sec (20 W/min). The selection of the appropriate protocol per individual participant was based on expert opinion of the test assistant. Criteria to stop the test were volitional exhaustion or failure in keeping a constant cadence above the preset value. PO (W) was measured during the test. POpeak was defined as the highest PO attained during this specific synchronous GXT. For the one-minute protocol POpeak (W) was defined as the highest PO that was maintained for at least 30 seconds. For the ramp protocol the highest PO achieved during the test was considered POpeak. Apart from PO, gas exchange was measured using the Cosmed (Cosmed, Roma, Italy), Cortex (Cortex, CORTEX Biophysik GmbH, Germany) or Oxycon (Erich Jaeger, Viasys Healthcare, Germany). The equipment was calibrated before each test. VO_2peak (l/min) and the peak respiratory exchange ratio (RERpeak) were defined as the highest 30-second average for VO_2 (l/min) and RER, respectively. HRpeak (bpm) was defined as the highest heart rate achieved during the test.

Determinants

During the medical screening, age (years), sex, height (m), TSI (years), lesion level, completeness of the lesion (using the ASIA Impairment Scale (AIS, category A, B, C, D) (39)) and average handcycling weekly training hours in the last 3 months (hours) were obtained anamnestically. As all individual lesion levels would create too many dummy variables for the analyses, and only 12 individuals with a tetraplegia (of 128 participants) could be included, lesion level was split in two categories: (1) above Th6 and (2) equal to or below Th6 to investigate the effect of loss of sympathetic cardiac innervation (lesion level above Th6) and preserved sympathetic cardiac innervation (lesion level equal to or below Th6) on POpeak (40). Body mass (kg) was measured on a wheelchair scale with the wheelchair included. Afterwards the mass of the wheelchair was weighted separately and subtracted from the total mass to obtain the body mass of the participant. Body Mass Index (BMI, in kg/m^2) was calculated by dividing the body mass by the squared height. Waist circumference (cm) was measured three times at the level of the umbilicus in supine position. The average of the three measurements was used for analysis. Handcycling classification was determined by an UCI certified Paracycling classifier, following the UCI

Para-cycling Regulations: ranging from H1 to H5, in which H1 is the most impaired class and H5 the least impaired class (41).

Statistical Analysis

The analyses were performed using SPSS (IBM SPSS Statistics 20, SPSS, Inc, Chicago, IL, USA) and MLWin software (42).

Descriptives

Means and standard deviations (M ± SD) were calculated for outcome measures and determinants, and data was tested for normality by means of the Kolmogorov–Smirnov test with Lilliefors Significance Correction and the Shapiro–Wilk test. In addition, z-scores for skewness and kurtosis were calculated.

Splitting the data

In order to validate the models, the group of 128 participants was randomly split into two samples, using random sample of cases in SPSS: (1) one sample to develop the predicted models (80% of the data; model group) and (2) one sample to cross-validate the models (20% of the data; validation group). This is based on the statement that the ratio of number of independent variables to the number of participants should be at least 1:10 in a multiple linear regression analysis (43). In this study, ten possible independent variables were identified; therefore, around 100 participants deemed necessary for the development of the model. First, the two sample groups were checked for systematic differences in baseline values to ensure equality between groups. Thereafter, the predictive model was developed using a multi-level regression analysis to correct for rehabilitation center (i.e., to correct for possible differences in test setting / testers / protocols between the 11 rehabilitation centers). A two-level model was created with participant as first level and center as second level.

Outcome measures and determinants

The dependent variables of the analyses were POpeak (W) and POpeak/kg (W/kg). POpeak/kg was chosen to compare the results of the present study with previous literature (18), and because of the importance of values in W/kg for the HandbikeBattle population as they are participating in an uphill mountain race. The independent variables were: age (years), sex (0=male, 1=female), body mass (kg), BMI in kg/m^2, waist circumference (cm), TSI (years), lesion level (two categories: (1) above Th6 and (2) equal to or below Th6), handcycling classification (two categories: (1) H1-H3 and (2) H4-H5), completeness of the lesion (two categories: (1) motor complete (AIS A+B) and (2) motor incomplete (AIS C+D)) and average handcycling weekly training hours in the last 3 months (h).

Predictive models

First, all variables were checked for multicollinearity as described by Field (44). Thereafter, all applicable independent variables were used in each of the two theoretical models. For the two statistical models, first, a series of univariate regression models was used within the model group to determine significant associations per variable ($p<0.10$). Thereafter a multi-level regression analysis was performed with all significant variables from the univariate analysis, using a backward elimination technique to develop a model with significant variables only ($p<0.05$). Only simple main effects of determinants were evaluated. For all four models the proportion of explained variance (R^2) was calculated.

Validation of the models

With the use of the developed models, the estimated POpeak was calculated in the validation group (N=24). Thereafter, these estimated scores for POpeak were compared to the (actual) measured POpeak (N=24). Systematic differences between these values were investigated with the paired-samples t-test. The intraclass correlation coefficient was used to measure relative agreement (ICC, two-way random, absolute agreement, single measures) and Bland-Altman plots with 95% limits of agreement (LoA) to measure absolute agreement (45,46). The following interpretation was used for the ICC: < 0.40 “poor”, 0.40 - 0.59 “fair”, 0.60 - 0.74 “good”, ≥ 0.75 “excellent” (47).

Reference values

Reference values for POpeak, POpeak/kg, VO_2peak and VO_2peak/kg based on lesion level and sex were developed with the data of all 128 participants. Quintiles were defined based on percentiles: Poor (below 20%), Fair (20% to 40%), Average (40% to 60%), Good (60% to 80%), and Excellent (above 80%), as described by Janssen et al (29).

Results

Descriptives

Means and standard deviations of outcome measures and determinants are depicted in table 1. Main outcome measures were normally distributed.

Splitting the data

No systematic differences in personal and fitness characteristics were observed between the model group and validation group (table 1).

Predictive models

For both models of POpeak and POpeak/kg, a two-level model was created with participant as first level and center as second level. For both models the -2log likelihood did not significantly change after adding center as a level to the constant, i.e. rehabilitation center did not have a substantial effect on the outcome.

Of the possible determinants, lesion level and handcycling classification showed a significant correlation ($r = 0.46$, $p < 0.001$, tolerance = 0.79, variance inflation factor (VIF) = 1.27). Body mass, BMI and waist circumference showed a significant correlation as well ($r \geq 0.78$, $p < 0.001$, tolerance ≤ 0.33, VIF ≥ 3.07 for all correlations). This indicates multicollinearity and, therefore, these variables were not analyzed in combination with each other in the models. Separate models were developed for these variables: BMI and lesion level were used as determinants in the final four models based on significance and proportion of explained variance.

Theory-driven models

In the theoretical model for POpeak, sex, lesion level, handcycling training hours and age were significant determinants. In the theoretical model for POpeak/kg, sex, lesion level, handcycling training hours, BMI and age were significant determinants. R^2 was 42% for both models (table 2).

Statistically-driven models

In the statistical model for POpeak, sex, lesion level, handcycling classification, body mass, BMI and handcycling training hours were significant determinants based on the univariate analysis. In the backward analysis sex, lesion level and handcycling training hours remained significant and formed the final statistical model for POpeak (R^2 = 39%) (table 2).

In the statistical model for POpeak/kg, age, lesion level, body mass, BMI, waist circumference and handcycling training hours were significant determinants based on the univariate analysis. In the backward analysis, lesion level, handcycling training hours and BMI remained significant and formed the final statistical model for POpeak/kg (R^2 = 30%) (table 2).

Validation of the models

For all four models, no systematic differences were found between the predicted POpeak and the measured POpeak. Validation of the models showed varying results, depending on the model (table 3). A fair relative agreement (ICC = 0.43) for the theoretical POpeak model was found, while the Bland-Altman plot showed a large variation (95% LoA -69 – 54 W) indicating a low absolute agreement (figure 1A). The theoretical POpeak/kg model showed a good relative agreement (ICC = 0.60), however, the Bland-Altman plot showed a large variation (95% LoA -0.78 – 0.57 W/kg) for this model as well (figure 1B). A poor relative agreement (ICC = 0.35) for the statistical POpeak model was found, which was supported by the large variation observed in the Bland-Altman plot (95% LoA -64 – 57 W) (figure 1C). Lastly, the statistical POpeak/kg model showed a fair relative agreement (ICC = 0.43), with a large variation (95% LoA -0.92 – 0.68 W/kg) in the Bland-Altman plot (figure 1D).

Reference values

Table 4 and table 5 show reference values for POpeak, POpeak/kg, VO_2peak and VO_2peak/kg based on lesion level and sex, developed with the data of all 128 participants.

Discussion

This study is the first to have developed and validated predictive models and reference values for synchronous handcycling. Four predictive models on POpeak (W and W/kg) were developed in a group of recreational handcyclists: two theory-driven models and two statistically-driven models. The theoretical models showed a somewhat higher explained variance than the statistical models, although overall the explained variance was low for all four models (R^2 ranged from 30% to 42%). Validation of the models showed a poor to good relative agreement, depending on the model, with a low absolute agreement for all models. In accordance with the third aim, reference values for POpeak, POpeak/kg, VO_2peak and VO_2peak/kg based on lesion level and sex were developed.

Predictive models

Due to missing data, both theoretical models were based on fewer participants (N=84) than the statistical models (N=94-95) (table 2). However, these models showed more statistically significant determinants and a higher explained variance than the statistical models. This might be due to a different interdependent association between the determinants in these models; in the theoretical models all determinants were included simultaneously (forced entry) based on our understanding of interdependency, whereas in the statistical models first an univariate analysis was performed. In this univariate analysis, some determinants were excluded from the model based on their individual association with POpeak, obviously without considering their possible indirect association with POpeak through their interactions with other determinants. Compared to theory-driven modeling, this is a disadvantage of stepwise statistical modeling as only mathematical criteria are used to select determinants (44). In future studies, it could be interesting to focus on these possible interactions between determinants when modeling physical capacity in individuals with SCI.

Theory-driven models

In this study, two theory-driven models for POpeak were developed using multi-level regression analysis. The selection of determinants was based on theoretical constructs, investigated in previous wheelchair and arm ergometry literature concerning individuals with a SCI. The aim was to gain more insight in the underlying determinants influencing physical capacity in individuals with SCI during handcycling. The results showed that sex, lesion level, handcycling training hours and age are significant determinants for POpeak (table 2). Of these determinants handcycling training hours is the only determinant that can be influenced. Therefore, in order to increase physical capacity in individuals with a SCI during handcycling, individually optimized training intensity and volume should be encouraged. Another modifiable determinant, BMI, was positively related to POpeak, although not significant, and inversely related to POpeak/kg, which indicates a decrease in physical capacity with every increase in BMI. This can partly be explained by the shared term for mass in the outcome measure (POpeak/kg) and the determinant (BMI). Comparable relationships were previously described by Janssen et al. (29) and Simmons et al. (18) in wheelchair ergometry and asynchronous arm ergometry, respectively. They explain that an elevated BMI in this population is, therefore, probably related to overweight due to adipose tissue and a low physical activity, instead of a large muscle mass. BMI was chosen in this study (instead of bio impedance analysis or DXA) due to its wide use in literature and clinical practice, inexpensiveness, applicability, and in order to compare our results with previous literature about predictive models in wheelchair exercise and asynchronous arm ergometry.

Statistically-driven models

Next to the theory-driven models, two statistically-driven models were developed. The aim was to use multi-level regression analyses with a backward elimination technique to accurately predict POpeak during handcycling based on statistically significant determinants. Results showed that only three determinants appeared to be statistically significant determinants (sex, lesion level and handcycling training for POpeak, and lesion level, handcycling training and BMI for POpeak/kg) following the current statistical selection criteria and backward approach. In previous literature, only statistical models were developed to investigate the association between POpeak and participant characteristics, based on wheelchair testing and asynchronous arm ergometry. Simmons et al. (18) developed a model for POpeak during asynchronous arm ergometry in untrained individuals with a SCI based on motor level of injury, functional classification and sex (R^2 = 0.57) and a model for POpeak/kg based on functional classification, BMI and motor level of injury (R^2 = 0.57) using (forward) stepwise regression. Other possible factors such as age, TSI and completeness were not significantly correlated to POpeak in the study of Simmons et al.,

(18) comparable to the results in the present study. An important difference between the study by Simmons et al. and the present study is the determinant handcycling training (hours). This determinant was significant in both statistical models in the present study, however, was not analyzed in the study by Simmons et al. Janssen et al. (29) found a comparable determinant, activity level, to be significantly related to POpeak in wheelchair exercise testing. Moreover, several other studies highlighted the relationship between activity level or sports participation and physical capacity in individuals with a SCI during wheelchair testing (32,35) and asynchronous arm ergometry (9,17).

Despite the significant determinants that were found, a large part of the variance in the present study remained unexplained (58-70%). This might have several reasons. First, due to the multicenter character of the study, different test assistants performed the tests and different test equipment and protocols were used. This causes inevitable variability in test results. Although, in the present study, no significant differences were found between rehabilitation centers, test equipment and protocols, it would be optimal to standardize these measures in order to pursue homogeneity. However, the reader should be aware that in order to achieve a large number of participants in rehabilitation related research, homogeneity is only possible to a certain extent. In this study, a correction was made for the possible (non-significant) differences between rehabilitation centers by multi-level regression analysis. Second, we need to critically evaluate the way determinants are reported and consider other possible determinants. For example, handcycling training was reported; however, other activities of daily living and lifestyle factors as well as other types of training (e.g. swimming, wheelchair rugby, but also strength training) were not taken into account as the response rate on these separate questions and the validity of the answers were considered too low to be representative. This is unfortunate, as these factors might explain a larger part of the variance than handcycling training alone. Moreover, training hours do not take the actual intensity level into account. Therefore, an overall, easy to use measure of training load should be considered such as Training Impulse based on session ratings of perceived exertion (sRPE) (48,49), to increase the proportion of explained variance.

As emphasized by Van Der Woude et al. (50), POpeak is associated with several factors, including the factors that were taken into account in the present study. POpeak is, however, also directly related to the mode of exercise (e.g. handrim wheelchair or handbike propulsion), including notions of efficiency, skill and talent, as well as aerobic exercise (cardiorespiratory) and anaerobic capacity. POpeak is, therefore, a general measure of handcycling physical capacity. This is in contrast to VO_2peak, as VO_2peak is a general measure of cardiorespiratory function only (50,51). Therefore, more factors associated with POpeak should be taken into account. For example, in a previous study by Janssen et al. (30) a strong association was found between anaerobic POpeak and aerobic POpeak (R^2 = 81%) in individuals with a SCI on a wheelchair ergometer. Future studies could focus on this association in handcycling with, for example, a Wingate Test, which might lead to a higher explained variance and, subsequently, better estimation of POpeak.

Validation of the models

To the authors' knowledge, this is the first study that investigated validity of a POpeak prediction model in arm exercise. Despite a good relative agreement for the theoretical POpeak/kg model, all models showed a low absolute agreement as represented by the high variation in the Bland-Altman plots (figure 1). Although a high relative and absolute agreement are desirable, it must be emphasized that these models were not designed to replace the GXT. It is, therefore, not necessarily needed to predict the exact POpeak, a certain valid range, however, is a prerequisite. It has been suggested that a test duration of 8 – 12 minutes would be optimal to achieve peak physiological responses during a GXT (21,25), although it is important to mention that the optimal test duration for arm exercise is not known (52). This test duration is important, as it is inherent to the number of steps and the step size of the protocol. Studies have shown that when the step size is too large, and consequently the test is too short, peak physical capacity tends to be overestimated and studying the effect of certain therapy or training is less reliable (25). However, when the test is too long due to the small step size or long step duration, peak physical capacity tends to be underestimated (21,24). As an average test duration of 10 minutes ± 20% is said to be optimal, it could be argued that a predicted POpeak within a range of ± 20% is a valid value to use in the selection of an individualized GXT protocol. In this study, depending on the model, 52 – 67% of the predicted POpeak values fell within this range. This indicates that the validity of the models is not high enough to solely base GXT protocol selection on. Therefore, future research should focus on improving the validity of these models and diminishing the large proportion of unexplained variance.

Reference values

To date, this is the first study that describes reference values for (synchronous) handcycling based on a large group of handcycle users with SCI. Comparing the results to previous literature, it has to be emphasized that our group was heterogeneous and that not all participants were completely untrained. In the study by Lovell et al. (53), a mean POpeak of 121 W was found for untrained handcyclists with paraplegia, which is comparable to the results in the present study (120 – 136W). It must be emphasized that it is unclear whether synchronous or asynchronous arm cranking was performed in the study by Lovell et al. Due to the heterogeneity of the population in the present study, the reference values will give a good reflection of the diversity in the SCI population. However, individuals with a very low physical capacity or absent training motivation are probably not represented in this study, as these individuals are not motivated to participate in a mountain race. Moreover, elite handcyclists did not participate in our study, as a POpeak of 210 W as described by Lovell et al. (53) for "trained" handcyclists with a SCI was reached by none of the participants in the present study. This has to be considered when interpreting the predictive models and reference values.

Next to training status, other factors need to be kept in mind comparing the results of the present study to previous research. For example, test device (wheelchair ergometry versus arm ergometry versus handcycling), propulsion mode (asynchronous versus synchronous), test protocol and other participant characteristics. Overall, the reference values of the present study were higher compared to values found in previous studies focusing on asynchronous arm ergometry. Simmons et al. (18) found an average POpeak of 62 – 78 W and 0.85 – 0.98 W/kg during (asynchronous) arm ergometry for men with paraplegia, compared to 120 – 136 W and 1.52 – 1.70 W/kg, respectively, for the group with low paraplegia in the present study. Next to POpeak, VO_2peak showed higher values in the present study: Simmons et al. (18) found an average VO_2peak of 1.28 – 1.41 L/min and 15.31 – 17.69 mL/kg/min during arm ergometry for men with paraplegia, compared to an average VO_2peak of 1.95 – 2.20 L/min and 24.61 – 27.42 mL/kg/min, respectively, in the present study. Earlier reviews by Haisma et al. (7) and Valent et al. (54) studying reference values for individuals with paraplegia during asynchronous arm ergometry support the finding of Simmons et al. The reviews showed a POpeak of 66 – 117 W (7) and a VO_2peak of 1.06 – 2.34 L/min (7) and 1.33 – 1.90 L/min (54).

The reference values found in the present study are comparable to a previous study investigating synchronous handcycling (55). Janssen et al. performed a descriptive study with 16 male handcycle users, measuring physical capacity by means of a GXT in an add-on handcycle on a treadmill (55). Although not exclusively individuals with a SCI were studied, they found similar values for the group with lower-limb disabilities: 129 ± 26 W and 1.64 ± 0.32 W/kg, comparable to results of the present study. Dallmeijer et al. (3) studied physical capacity by means of a GXT in an add-on handcycle on a treadmill in 9 men with a paraplegia and found a POpeak of 117 ± 32 W and a VO_2peak of 1.88 ± 0.44 L/min. These results are slightly lower than in the present study.

Implications

The theoretical POpeak/kg model was the best predicting model to assess POpeak, with an explained variance of 42% and ICC of 0.60. However, a large part of the variance still remained unexplained and the Bland-Altmann plot showed a low absolute agreement. Moreover, the finding that only 67% of the predicted POpeak values fell into the range of ± 20% indicates that the validity of this model is not high enough to solely base GXT protocol selection on. Therefore, the models should be used with caution and only in addition to expert opinion of the practitioner when there is indecisiveness in what protocol to choose. It must be explicitly emphasized that the models should not be used to replace a GXT. In future studies standardization of test setting and protocol is necessary.

The same large part of unexplained variance is reflected on the reference values. Nevertheless, this is the first study to describe reference values for (synchronous) handcycling in individuals with a SCI. Although the values should be used with caution, they give a global overview of the physical capacity of individuals with a SCI during handcycling. As these values are based on a large heterogeneous group, they give an indication of the normal variation in the SCI population, for both men and women, and only applicable to synchronous handcycling.

Study limitations

There was variation in the measurement set-up due to the fact that tests were performed in 11 different rehabilitation centers. Although, in the present study, no significant effect of rehabilitation centers was found, it would be optimal to standardize these measures in order to pursue homogeneity. Secondly, due to the low number of individuals with a tetraplegia (N=12), it was not possible to divide the group in people with tetraplegia and paraplegia. The results of this study are, therefore, not applicable to individuals with a tetraplegia. Moreover, due to the relatively low number of female participants (N=22) it was not possible to define reference values based on sex and lesion level together. Therefore, separate reference values were defined; 1) for lesion level, and 2) for sex. Lastly, possible important determinants such as training load were not taken into account. This might be interesting for future research.

Conclusion

This study is the first to have developed and validated predictive models and reference values for synchronous handcycling. Lesion level, handcycling training hours and sex or BMI appeared to be significant determinants of POpeak in handcyclists with SCI in all four models. The theoretical models showed the highest proportion of explained variance. Validation showed varying relative agreement, and a low absolute agreement. Moreover, a large part of the variance remained unexplained in all models. Therefore, these models and reference values might be useful in clinical practice, but should not replace a GXT. Both models and reference values provide insight in physical capacity of the diverse SCI population, based on a relatively large sample performing synchronous handcycling GXT.

Implications for Rehabilitation

- Individualization of the graded exercise test protocol is very important to attain the true peak physical capacity in individuals with spinal cord injury.
- The main determinants to predict peak power output during a handcycling graded exercise test for individuals with a spinal cord injury are lesion level, handcycling training hours and sex or body mass index.
- The predictive models for peak power output should be used with caution and should not replace a graded exercise test.

Declaration of Conflicting Interests

The Authors declare that there is no conflict of interest.

Acknowledgements

*HandbikeBattle group name: Paul Grandjean Perrenod Comtesse, Adelante Zorggroep, Hoensbroek, The Netherlands. Eric Helmantel, University Medical Center Groningen, Center for Rehabilitation Beatrixoord, Groningen, The Netherlands. Mark van de Mijll Dekker, Heliomare Rehabilitation Center, Wijk aan Zee, The Netherlands. Maremka Zwinkels, Rehabilitation Center De Hoogstraat, Utrecht, The Netherlands. Misha Metsaars, Libra Rehabilitation and Audiology, Eindhoven, The Netherlands. Lise Wilders, Sint Maartenskliniek, Nijmegen, The Netherlands. Linda van Vliet, Amsterdam Rehabilitation Research Center | Reade, Amsterdam, The Netherlands. Karin Postma, Rijndam Rehabilitation Center, Rotterdam. Bram van Gemeren, Roessingh Rehabilitation Center, Enschede, The Netherlands. Selma Overbeek, Jeroen Bosch Hospital, Tolbrug Rehabilitation Centre, ‘s-Hertogenbosch, The Netherlands. Alinda Gjaltema, Vogellanden, Zwolle, The Netherlands.

References

1. Abel T, Vanlandewijck Y, Verellen J. No Title. In: Goosey-Tolfrey V, editor. Wheelchair Sport. Human Kinetics; 2010. p. 187–97.

2. Arnet U, Van Drongelen S, Scheel-Sailer A, et al. Shoulder load during synchronous handcycling and handrim wheelchair propulsion in persons with paraplegia. J Rehabil Med. 2012;44(3):222–8.

3. Dallmeijer AJ, Zentgraaff IDB, Zijp NI, et al. Submaximal physical strain and peak performance in handcycling versus handrim wheelchair propulsion. Spinal Cord. 2004;42(2):91–8.

4. Van Der Woude L, Dallmeijer AJ, Janssen TWJ, et al. Alternative modes of manual wheelchair ambulation: An overview. Am J Phys Med Rehabil. 2001; (October):765-77.

5. Valent LJM, Dallmeijer AJ, Houdijk H, et al. Effects of Hand Cycle Training on Physical Capacity in Individuals With Tetraplegia: A Clinical Trial. Phys Ther. 2009;89(10):1051–60.

6. Van Den Berg-Emons RJ, Bussmann JB, Haisma JA, et al. A Prospective Study on Physical Activity Levels After Spinal Cord Injury During Inpatient Rehabilitation and the Year After Discharge. Arch Phys Med Rehabil. 2008;89(11):2094–101.

7. Haisma JA, Van Der Woude LH V, Stam HJ, et al. Physical capacity in wheelchair-dependent persons with a spinal cord injury: A critical review of the literature. Spinal Cord. 2006;44(11):642–52.

8. Jacobs PL, Nash MS. Exercise recommendations for individuals with spinal cord injury. Sports Med. 2004;34(11):727–51.

9. Muraki S, Tsunawake N, Tahara Y, et al. Multivariate analysis of factors influencing physical work capacity in wheelchair-dependent paraplegics with spinal cord injury. Eur J Appl Physiol Occup Physiol. 2000;81(1–2):28–32.

10. De Groot S, Dallmeijer AJ, Post MWM, et al. The longitudinal relationship between lipid profile and physical capacity in persons with a recent spinal cord injury. Spinal Cord. 2008;46(5):344–51.

11. van Koppenhagen CF, Post M, de Groot S, et al. Longitudinal relationship between wheelchair exercise capacity and life satisfaction in patients with spinal cord injury: A cohort study in the Netherlands. J Spinal Cord Med. 2014;37(3):328–37.

12. Manns PJ, Chad KE. Determining the relation between quality of life, handicap, fitness, and physical activity for persons with spinal cord injury. Arch Phys Med Rehabil. 1999;80(12):1566–71.

13. Van Velzen JM, de Groot S, Post MWM, et al. Return to Work After Spinal Cord Injury. Am J Phys Med Rehabil. 2009;88(1):47–56.

14. Noreau L, Shephard RJ. Return to work after spinal cord injury: The potential contribution of physical fitness. Paraplegia. 1992;30(8):563–72.

15. Goosey-Tolfrey VL, Sindall P. The effects of arm crank strategy on physiological responses and mechanical efficiency during submaximal exercise. J Sports Sci. 2007;25(4):453–60.

16. Hooker SP, Wells CL. Aerobic power of competitive paraplegic road racers. Paraplegia. 1992;30(6):428–36.

17. Hutzler Y, Ochana S, Bolotin R, et al. Aerobic and anaerobic arm-cranking power outputs of males with lower limb impairments: relationship with sport participation intensity, age, impairment and functional classification. Spinal Cord. 1998;36(3):205–12.

18. Simmons OL, Kressler J, Nash MS. Reference fitness values in the untrained spinal cord injury population. Arch Phys Med Rehabil. 2014;95(12):2272–8.

19. Mossberg K, Willman C, Topor MA, et al. Comparison of asynchronous versus synchronous arm crank ergometry. 1999;569–74.

20. De Groot S, Postma K, Van Vliet L, et al. Mountain time trial in handcycling: Exercise intensity and predictors of race time in people with spinal cord injury. Spinal Cord. 2014;52(6):455–61.

21. Buchfuhrer MJ, Hansen JE, Robinson TE, et al. Optimizing the exercise protocol for cardiopulmonary assessment. J Appl Physiol Respir Environ Exerc Physiol. 1983;55(5):1558–64.

22. Smith PM, Amaral I, Doherty M, et al. The influence of ramp rate on V̇O2peak and “excess” V̇O2 during arm crank ergometry. Int J Sports Med. 2006;27(8):610–6.

23. Eerden S, Dekker R, Hettinga FJ. Maximal and submaximal aerobic tests for wheelchair-dependent persons with spinal cord injury: a systematic review to summarize and identify useful applications for clinical rehabilitation. Disabil Rehabil.

2018;40(5):497–521.

24. Bentley DJ, Newell J, Bishop D. Incremental exercise test design and analysis: Implications for performance diagnostics in endurance athletes. Sport Med. 2007;37(7):575–86.

25. Myers J, Bellin D. Ramp exercise protocols for clinical and cardiopulmonary exercise testing. Sports Med. 2000;30(1):23–9.

26. Dallmeijer AJ, Kilkens OJE, Post MWM, et al. Hand-rim wheelchair propulsion capacity during rehabilitation of persons with spinal cord injury. J Rehabil Res Dev. 2005;42(3 Suppl 1):55–63.

27. Falkel JE, Sawka MN, Levine L, et al. Upper-body exercise performance: Comparison between women and men. Ergonomics. 1986;29(1):145–54.

28. Coutts KD, Rhodes EC, McKenzie DC. Maximal exercise responses of tetraplegics and paraplegics. J Appl Physiol. 1983;55(2):479–82.

29. Janssen T, Dallmeijer AJ, Veeger HEJ, et al. Normative values and determinants of physical capacity in individuals with spinal cord injury. JRRD. 2002;39(1). p. 29–39.

30. Janssen TW, Van Oers CA, Hollander AP, et al. Isometric strength, sprint power, and aerobic power in individuals with a spinal cord injury. Med Sci Sports Exerc. 1993;25:863–70.

31. Sawka MN, Glaser RM, Laubach LL, et al. Wheelchair exercise performance of the young, middle-aged, and elderly. J Appl Physiol. 1981;50(4):824–8.

32. Dallmeijer AJ, Hopman MT, Van As HH, et al. Physical capacity and physical strain

in persons with tetraplegia; the role of sport activity. Spinal Cord. 1996;34(11):1173–6.

33. Hicks AL, Martin Ginis KA, Pelletier CA, et al. The effects of exercise training on physical capacity, strength, body composition and functional performance among adults with spinal cord injury: A systematic review. Spinal Cord. 2011;49(11):1103–27.

34. Janssen TW, Van Oers CA, Rozendaal EP, et al Changes in physical strain and physical capacity in men with spinal cord injuries. Med Sci Sports Exerc. 1996;28(5):551–9.

35. Dallmeijer AJ, Van Der Woude LHV, Hollader PA, et al. Physical performance in persons with spinal cord injuries after discharge from rehabilitation. Med Sci Sport Exerc. 1999;31(8):1111–1117 7p.

36. De Groot S, Van Der Scheer JW, Bakkum AJT, et al. Wheelchair-specific fitness of persons with a long-term spinal cord injury: cross-sectional study on effects of time since injury and physical activity level. Disabil Rehabil. 2016;38(12):1180–6.

37. Hoekstra S, Valent L, Janssen T, et al. Transferability of exercise test results among three arm ergometers. Poster session presented at: Biannual conference of the International Paralympic Committee, VISTA; 2015; Gerona Spain.

38. Smith PM, Doherty M, Drake D, et al. The influence of step and ramp type protocols on the attainment of peak physiological responses during arm crank ergometry. Int J Sports Med. 2004;25(8):616–21.

39. Kirshblum SC, Burns SP, Biering-Sorensen F, et al. International standards for neurological classification of spinal cord injury (Revised 2011). J Spinal Cord Med. 2011;34(6):535–46.

40. Krassioukov A, West C. The Role of Autonomic Function on Sport Performance in Athletes With Spinal Cord Injury. PM&R. 2014;6(8):S58–65.

41. UCI Cycling Regulations, part 16 Para-cycling. 2017. p. 1–75. Available from: http://www.uci.ch/inside-uci/rules-and-regulations/regulations/

42. Rasbash J, Charlton C, Browne W, et al. MLwiN Version 2.02. Centre for Multilevel Modelling, University of Bristol; 2005.

43. Altman DG, Royston P. What do we mean by validating a prognistic model? Stat Med. 2000;19:453–73.

44. Field A. Discovering Statistics using IBM SPSS Statistics. 4th ed. Carmichael M, editor. SAGE; 2013. 321-326 p.

45. Bland JM, Altman DG. Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement. Lancet. 1986;327:307–10.

46. Mantha S, Roizen MF, Fleisher L et al. Comparing methods of clinical measurement: reporting standards for bland and altman analysis. Anesth Analg. 2000;90(3):593–602.

47. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess. 1994;6(4):284–90.

48. Foster C, Florhaug JA, Franklin J, et al. A New Approach to Monitoring Exercise Training. J Strength Cond Res. 2001;15(1):109–15.

49. De Groot S, Hoekstra SP, Grandjean Perrenod Comtesse P, et al. Relationships between internal and external handcycle training load in people with spinal cord injury training for the HandbikeBattle. J Rehabil Med. 2018, 50(3), 261-268

50. Van Der Woude LHV, De Groot S, Van Drongelen S, et al. Evaluation of Manual Wheelchair Performance in Everyday Life. Top Spinal Cord Inj Rehabil. 2009;15(2):1-15.

51. Kilkens OJ, Dallmeijer AJ, Nene AV, et al. The longitudinal relation between physical capacity and wheelchair skill performance during inpatient rehabilitation of people with spinal cord injury. Arch Phys Med Rehabil. 2005;86(8):1575–81.

52. Maher JL, Cowan RE. Comparison of 1- Versus 3-Minute Stage Duration During Arm Ergometry in Individuals With Spinal Cord Injury. Arch Phys Med Rehabil. 2016 Nov;97(11):1895–900.

53. Lovell D, Shields D, Beck B, et al. The aerobic performance of trained and untrained handcyclists with spinal cord injury. Eur J Appl Physiol. 2012;112(9):3431–7.

54. Valent L, Dallmeijer A, Houdijk H, et al. The effects of upper body exercise on the physical capacity of people with a spinal cord injury: a systematic review. Clin Rehabil. 2007;21(4):315–30.

55. Janssen TW, Dallmeijer AJ, van der Woude LH. Physical capacity and race performance of handcycle users. J Rehabil Res Dev. 2001;38(1):33–40.

Table 1. Participant characteristics of the total group (N=128), the model group (80% of data, N=104), and the validation group (20% of data, N=24).

	Total group (N = 128)		**Model group** (N = 104)	**Validation Group** (N = 24)
	M ± SD or N	*N total*	*M ± SD or N*	*M ± SD or N*
SCI/spina bifida	118/10	128	96/8	22/2
Lesion level (>Th6/≤Th6)	37/86	123	32/68	5/18
Completeness (motor compl/incompl)	77/41	118	61/35	16/6
Sex (male/female)	106/22	128	85/19	21/3
Age (years)	39 ± 12	128	39 ± 12	39 ± 12
TSI (years)	10 ± 10	119	10 ± 10	10 ± 9
Height (m)	1.80 ± 0.10	127	1.79 ± 0.10	1.80 ± 0.11
Body Mass (kg)	78 ± 17	127	78 ± 16	79 ± 18
BMI (kg/m^2)	24 ± 4	126	24 ± 4	24 ± 4
Waist circumference (cm)	91 ± 15	116	91 ± 15	88 ± 17
Handcycling training (h)	3.39 ± 3.70	121	3.51 ± 3.84	2.84 ± 2.99
Handcycling classification (H1-H3/H4-H5)	67/57	124	55/46	12/11
POpeak (W)	119 ± 34	128	119 ± 33	121 ± 40
POpeak/kg (W/kg)	1.54 ± 0.47	127	1.54 ± 0.46	1.56 ± 0.51
VO_2peak (L/min)	1.91 ± 0.58	126	1.88 ± 0.56	2.05 ± 0.66
VO_2peak/kg (mL/kg/min)	24.93 ± 7.91	125	24.58 ± 7.60	26.51 ± 9.17
HRpeak (bpm)	171 ± 22	124	171 ± 22	174 ± 23
RERpeak	1.21 ± 0.12	115	1.21 ± 0.12	1.22 ± 0.11
Cyclus 2 / Tacx / Arm ergometer	35/24/69	128	29/22/53	6/2/16
1 min / ramp	79/49	128	66/38	13/11

SCI: spinal cord injury, TSI: time since injury, BMI: Body Mass Index, POpeak: peak power ouput, VO_2peak: peak oxygen uptake, HRpeak: peak heart rate, RERpeak: peak respiratory exchange ratio. Lesion level: two categories: (1) above Th6 and (2) equal to or below Th6. Completeness: AIS (two categories: (1) motor complete (AIS A+B) and (2) motor incomplete (AIS C+D)), handcycling training: average handcycling weekly training hours in the last 3 months, handcycling classification: two categories: (1) H1-H3 and (2) H4-H5. Measurement device: cyclus 2, Tacx or arm ergometer. Protocoltype: 1 minute step protocol or ramp protocol.

Table 2. Results for both theoretical models (with all potential determinants) and for both statistical models (after backward regression analyses) to predict absolute and relative POpeak.

	Theoretical models						Statistical models					
	POpeak (N=84)			*POpeak/kg (N=84)*			*POpeak (N=95)*			*POpeak/kg (N=94)*		
	β (SE)	*95%CI*	*p*-value	β (SE)	*95%CI*	*p*-value	β (SE)	*95%CI*	*p*-value	β (SE)	*95%CI*	*p*-value
Intercept	107.05 (18.54)	70.7 143.4	< 0.01	2.94 (0.26)	2.43 3.44	< 0.01	99.97 (5.14)	89.9 110.0	< 0.01	2.36 (0.23)	1.91 2.81	< 0.01
Sex	-41.13 (7.88)	-56.6 -25.7	< 0.01	-0.38 (0.11)	-0.60 -0.16	< 0.01	-41.29 (6.96)	-54.9 -27.6	< 0.01	ns	NA	NA
Lesion level	26.67 (5.90)	15.1 38.2	< 0.01	0.33 (0.08)	0.17 0.49	< 0.01	28.88 (5.69)	17.7 40.0	< 0.01	0.31 (0.09)	0.13 0.49	< 0.01
Handcycling training (h)	1.82 (0.75)	0.35 3.29	0.02	0.03 (0.01)	-0.01 0.05	< 0.01	1.77 (0.71)	0.38 3.16	0.01	0.03 (0.01)	0.01 0.05	0.01
BMI (kg/m^2)	0.52 (0.84)	-1.13 2.17	0.54	-0.06 (0.01)	-0.08 -0.04	< 0.01	ns	NA	NA	-0.05 (0.01)	-0.07 -0.03	< 0.01
TSI (years)	0.18 (0.33)	-0.47 0.83	0.59	0.01 (0.01)	-0.01 0.03	0.23	ns	NA	NA	ns	NA	NA
Completeness	10.92 (6.24)	-1.31 23.15	0.08	0.10 (0.09)	-0.08 0.28	0.24	ns	NA	NA	ns	NA	NA
Age (years)	-0.59 (0.30)	-1.18 -0.002	0.05	-0.01 (0.004)	-0.02 -0.002	< 0.01	ns	NA	NA	ns	NA	NA
R^2	42%			42%			39%			30%		

β (SE) = beta with standard error. 95%CI = 95% confidence interval. R^2 = proportion of explained variance. Independent variables: sex (0=male, 1=female), lesion level (two categories: (1) above Th6 and (2) equal to or below Th6), average handcycling weekly training hours in the last 3 months (hours), Body Mass Index (BMI) in kg/m2, time since injury (TSI, years), completeness following AIS (two categories: (1) motor complete (AIS A+B) and (2) motor incomplete (AIS C+D)), age (years). ns = non significant; NA = not applicable.

Table 3. Validation of the models. Results of comparison between measured and predicted POpeak with intraclass correlation coefficient (N=24).

	Measured *M ± SD*	**Theoretical model** *M ± SD*	**ICC (95% CI)**	**Statistical model** *M ± SD*	**ICC (95% CI)**
POpeak (W)	121 ± 40	123 ± 17	0.43 (-0.03-0.74)*	126 ± 14	0.35 (-0.09-0.68)
POpeak/kg (W/kg)	1.56 ± 0.51	1.50 ± 0.31	0.60 (0.21-0.82)*	1.52 ± 0.23	0.43 (0.01-0.72)*

M ± SD indicates mean ± standard deviation. 95% CI = 95% confidence interval. *indicates significant correlation ($p < 0.05$)

Table 4.

Reference values for POpeak, POpeak/kg, VO_2peak and VO_2peak/kg, for participants with (1) lesion level above Th6 (>Th6) and (2) equal to or below Th6 (≤Th6). Poor (<20%), Fair (20-40%), Average (40-60%), Good (60-80%) and Excellent (>80%) (N=128).

Variable	Level	n	Poor	Fair	Average	Good	Excellent
POpeak (W)	>Th6	37	< 63	63 - 96	96 - 117	117 – 137	> 137
	≤Th6	86	< 101	101 - 120	120 - 136	136 – 154	> 154
POpeak/kg (W/kg)	>Th6	37	< 0.81	0.81 – 1.16	1.16 – 1.47	1.47 – 1.79	> 1.79
	≤Th6	85	< 1.31	1.31 – 1.52	1.52 – 1.70	1.70 – 2.01	> 2.01
VO_2peak (L/min)	>Th6	37	< 1.11	1.11 – 1.47	1.47 – 1.72	1.72 – 2.02	> 2.02
	≤Th6	84	< 1.65	1.65 – 1.95	1.95 – 2.20	2.20 – 2.49	> 2.49
VO_2peak/kg (mL/kg/min)	>Th6	37	< 15.53	15.53 – 17.57	17.57 – 21.90	21.90 – 26.63	> 26.63
	≤Th6	83	< 21.18	21.18 – 24.61	24.61 – 27.42	27.42 – 31.58	> 31.58

Table 5.

Reference values for POpeak, POpeak/kg, VO_2peak and VO_2peak/kg, for male (M) and female (F) participants. Poor (<20%), Fair (20-40%), Average (40-60%), Good (60-80%) and Excellent (>80%) (N=128).

Variable	Sex	n	Poor	Fair	Average	Good	Excellent
POpeak (W)	M	106	< 104	104 - 120	120 - 135	135 - 150	> 150
	F	22	< 69	69 - 81	81 - 92	92 – 107	> 107
POpeak/kg (W/kg)	M	105	< 1.18	1.18 – 1.47	1.47 – 1.65	1.65 – 2.05	> 2.05
	F	22	< 1.10	1.10 – 1.32	1.32 – 1.53	1.53 – 1.64	> 1.64
VO_2peak (L/min)	M	105	< 1.53	1.53 – 1.80	1.80 – 2.08	2.08 – 2.43	> 2.43
	F	21	< 1.09	1.09 – 1.33	1.33 – 1.66	1.66 – 1.82	> 1.82
VO_2peak/kg (mL/kg/min)	M	104	< 18.08	18.08 – 22.68	22.68 – 26.69	26.69 – 30.76	> 30.76
	F	21	< 17.89	17.89 – 22.11	22.11 – 24.45	24.45 – 27.77	> 27.77

Figure captions

Figure 1. Bland-Altman plots representing the absolute agreement between the predicted POpeak and the measured POpeak. Solid line represents the mean, dotted lines represent mean ± 2SD (95% LoA). Each circle represents a participant of the validation group. A: The difference in POpeak between the POpeak predicted with the theoretical model and the measured POpeak. B: The difference in POpeak/kg between the POpeak/kg predicted with the theoretical model and the measured POpeak/kg. C: The difference in POpeak between the POpeak predicted with the statistical model and the measured POpeak. D: The difference in POpeak/kg between the POpeak/kg predicted with the statistical model and the measured POpeak/kg.